Co-evolution of segregation guide DNA motifs and the FtsK translocase in bacteria: identification of the atypical Lactococcus lactis KOPS motif

Published online 28 February 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 12* ***5535–5545***
doi:10.1093/nar/gks171

Co-evolution of segregation guide DNA motifs and the FtsK translocase in bacteria: identification of the atypical *Lactococcus lactis* KOPS motif

Sophie Nolivos[1,2], Fabrice Touzain[3], Carine Pages[1,2], Michele Coddeville[1,2], Philippe Rousseau[1,2], Meriem El Karoui[3,*], Pascal Le Bourgeois[1,2] and François Cornet[1,2,*]

[1]Laboratoire de Microbiologie et de Génétique Moléculaire, CNRS, [2]Université de Toulouse, Université Paul Sabatier, F-31000, Toulouse and [3]INRA, UMR 1319, Institut Micalis, F-78352, Jouy-en-Josas, France

Received January 4, 2012; Revised January 31, 2012; Accepted February 1, 2012

ABSTRACT

Bacteria use the global bipolarization of their chromosomes into replichores to control the dynamics and segregation of their genome during the cell cycle. This involves the control of protein activities by recognition of specific short DNA motifs whose orientation along the chromosome is highly skewed. The KOPS motifs act in chromosome segregation by orienting the activity of the FtsK DNA translocase towards the terminal replichore junction. KOPS motifs have been identified in γ-Proteobacteria and in *Bacillus subtilis* as closely related G-rich octamers. We have identified the KOPS motif of *Lactococcus lactis,* a model bacteria of the *Streptococcaceae* family harbouring a compact and low GC% genome. This motif, 5′-GA AGAAG-3, was predicted *in silico* using the occurrence and skew characteristics of known KOPS motifs. We show that it is specifically recognized by *L. lactis* FtsK *in vitro* and controls its activity *in vivo*. *L. lactis* KOPS is thus an A-rich heptamer motif. Our results show that KOPS-controlled chromosome segregation is conserved in *Streptococcaceae* but that KOPS may show important variation in sequence and length between bacterial families. This suggests that FtsK adapts to its host genome by selecting motifs with convenient occurrence frequencies and orientation skews to orient its activity.

INTRODUCTION

Bacterial chromosomes are large, usually circular, DNA molecules that replicate from a unique origin (*ori*) and in a bidirectional manner to the opposite termination region (*ter*). This replicative organization is accompanied by a global *ori-ter* polarization of chromosome sequences that now appears as the most general and conserved feature of bacterial genome organization and dynamics [(1,2) for reviews]. The term 'replichore' has been coined to account for this *ori-ter* polarization (3).

Replichores are characterized by an asymmetric base composition (the GC-skew), with the leading strands being richer in guanine than lagging strands (4), and by an orientation bias of numerous DNA motifs (5,6). Of these, two have been shown to have a biological function: the chi sites that protect chromosomal DNA against degradation and promote homologous recombination [(7) for review] and the KOPS motifs that act in chromosome segregation by controlling the activity of the FtsK protein (8,9). Both chi and KOPS are over-represented in genomes (i.e. their occurrence is significantly higher than expected by chance). Their enrichment on leading strands (here referred to as leading strand skew or skew) is also significant (8,10), which means that they are more skewed than expected even when taking into account the GC skew. Depending on the phylum, chi sites are recognized by analogous systems such as RecBCD (11) or AddAB/RexAB (12,13). Consistently, although their distribution properties among bacterial genomes are conserved, known chi sites vary in length and sequence in several Proteobacteria and Firmicutes (10,14,15).

*To whom correspondence should be addressed. Tel: +33 561 335 985; Fax: +33 561 335 886; Email: francois.Cornet@ibcg.biotoul.fr
Correspondence may also be addressed to Meriem El Karoui. Tel: +33 1 3465 2071; Fax: +33 1 3465 2065; Email: meriem.el_karoui@jouy.inra.fr
Present addresses:
Sophie Nolivos, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK.
Fabrice Touzain, Laboratoire de Génomique des Microorganismes, Université Pierre et Marie Curie/CNRS, UMR 7238, F-75006, Paris, France.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

In contrast to the RecBCD/AddAB systems, most bacteria possess an FtsK orthologues (16,17). In *Escherichia coli*, FtsK acts both in chromosome segregation and cell division and is thought to couple these during the cell cycle [for reviews (16,18)]. Its N-terminal domain, as part of the cell division apparatus, targets FtsK to the division septum. Its C-terminal domain, $FtsK_C$, is the most conserved part of the protein and forms a dsDNA-translocase of the AAA+ ATPase family (19,20). $FtsK_C$ acts in the terminal region of the chromosome (21) and controls the last steps of segregation including the removal of catenation links between sister chromosomes and the resolution of chromosome dimers (22,23). $FtsK_C$ assembles as a hexameric motor upon interaction with the DNA (24). This interaction may occur with non-specific DNA but is preferential with KOPS motifs, which orient translocation in the direction specified by KOPS (24–26). KOPS are recognized by a winged-helix DNA-binding domain located in the extreme C-terminal FtsKγ subdomain (27). The crystal structure of a KOPS motif bound to *E. coli* or *Pseudomonas aeruginosas* FtsKγ revealed that three FtsKγ subdomains of the six present in an $FtsK_C$ hexamer are involved in the recognition of a single KOPS motif (26). Once assembled onto the DNA, FtsK translocates towards the terminal junction of KOPS polarity, at which the *dif* recombination site lies and finally activates XerCD-mediated recombination between *dif* sites to resolve chromosome dimers (28,29).

Most FtsK orthologues contain a conserved FtsKγ subdomain, suggesting that the KOPS-mediated control of FtsK translocation is conserved (16). Few data are, however, available for conservation of the KOPS motif. The proposed consensus for *E. coli* KOPS, 5′-GGGNAGGG-3′ (8), contains the 5′-GGGCAGGG -3′ motif that is also recognized by the γ-Protebacteria *Vibrio cholerae* and *P. aeruginosas* FtsK homologues (26,30). SpoIIIE, an FtsK homolog of the Firmicute *Bacillus subtilis*, does not recognize this motif but the 5′-GAGAAGGG-3′ motif (the SRS motif), equivalent to KOPS in length and only slightly divergent in sequence (31). This sequence conservation between KOPS motifs in phylogenetically distant species may suggest that KOPS/SRS represent prototypical motifs with conserved function in a wide range of bacterial phyla. A global search for skewed octamers whose skew increases towards the terminal region (called Architecture Imparting Sequences, AIMS) was conducted in 40 bacterial genomes (6). The 5′-GGGCAGGG-3′ motif in *E. coli* and 5′-GAGAAGGG-3′ motif in *B. subtilis* responded to these criteria. However, whereas the 5′-GGGCAGGG-3′ displays AIMS characteristics in most Proteobacteria, no common motif was identified in Firmicutes, suggesting that different and/or more criteria are needed to predict KOPS or the KOPS motifs can diverge in sequence inside a bacterial phylum.

Lactococcus lactis is a mesophilic lactic acid bacteria extensively used in dairy and health applications. Due to its industrial importance, it serves as a model organism for genetic and biochemical studies of this group of micro-organisms. Phylogenetically, *L. lactis* constitute the first branch that separates the *Streptococcaceae* from other Firmicutes. We have previously shown that *Streptococcaceae* possess an atypical Xer system, the XerS/dif_{SL} system, that uses a single recombinase, XerS, instead of the two XerC and XerD recombinase of classical Xer systems, and a divergent *dif* site for chromosome dimer resolution (32). Despite this difference, resolution of chromosome dimers by XerS/dif_{SL} requires the chromosome translocation activity of FtsK (29,32). We now report that the *L. lactis* chromosome contains KOPS motifs that orient the activity of FtsK. This motif, 5′-G AAGAAG-3′, differs from previously reported KOPS motifs both in sequence and length.

MATERIALS AND METHODS

Strains and plasmids

Strains used were derived from *E. coli* K12 strain LN2666 [W1485 F$^-$ *leu thyA thi deoB or C supE rpsL* (StR)] (33). Strains carrying the *dif-lacI-dif* cassette flanked by KOPS motifs were previously described (29). Strains carrying the 5′-GAAGAAG-3′ were constructed in a similar manner. The *3γ* constructs were designed as genes encoding repeats of FtsKγ subdomains (from residue 1266 of *E. coli* FtsK and residue 693 of *L. lactis* FtsK) separated by a 14 glycine-rich flexible linker (Figure 2A) followed by GT residues (*Kpn*I restriction site) before the second, and HM residues (*Nde*I restriction site) before the third FtsKγ copy. These constructs were ordered from GenScript (Piscataway, NJ, USA). For protein production and purification, the *3γ* genes were inserted into plasmid pFSKB3X (GTP technology, Toulouse, France), creating His-FLAG-*3γ* fusion genes in plasmids pCL380 (His-FLAG-$3\gamma_{Ec}$) and pCL381 (His-FLAG-$3\gamma_{Ll}$). For *in vivo* expression, relevant genes were inserted into a pGB2 (34) derivative carrying *araC-araBADp* expression cassette, yielding plasmids pCL374 (His-FLAG-$3\gamma_{Ec}$) and pCL375 (His-FLAG-$3\gamma_{Ll}$). XerC was produced *in vivo* from plasmid pFC241 [pGB2-*araBADp-xerC*; (29)].

Purification of 3γ proteins

Escherichia coli strain BL21(DE3) carrying plasmid pCL380 or pCL381 was grown in L-broth at 42°C to OD_{600} = 0.6. IPTG (0.1 mM) were added to the medium and incubated culture at 25°C for 3 h. Cells were recovered by centrifugation resuspended in buffer [50 mM phosphate buffer pH 8, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mg/ml lysosyme, 230 mg/ml RNaseA and EDTA-free proteases inhibitor cocktail (Roche)] and sonicated, and the lysate was cleared by centrifugation. His-FLAG-tagged 3γ proteins were purified on two successive nickel resin columns (1 ml His-trap HP, GE Healthcare) followed by a gel filtration columns (High-load 16/60 Superdex 200, GE Healthcare). Purified proteins were stored at −80°C in buffer containing 50 mM Hepes (pH 7.8), 40 mM KCl and 0.5 mM EDTA glycerol 10%.

ITC experiments

ITC experiments were performed using a MicroCal ITC200 Isothermal Titration Calorimeter. Experiments were carried out by titrating 3γ protein (50 μM) with DNA fragment as indicated [28 injections of 1.5 μl DNA solution at 450 μM (Figure 4D) or 250 μM (Figure 4E and F) in 3 s with a spacing of 180 s]. The stoichiometry of binding was obtained by fitting the ITC titration curves to the 'one set of site' model, assuming that the binding events were equivalent in the case of multiple binding. The best fitting model curves with corresponding stoichiometry are shown in Figure 4D and E.

EMSA experiments

Oligonucleotides were 5′ end-labelled using [γ-^{32}P] ATP and T4 DNA polynucleotide kinase and purified on MicroSpin G-25 column (GE Healthcare). DNA substrates were then prepared by hybridization of complementary labelled and unlabelled oligonucleotides. After 10 min denaturation in boiling water, the mixture was left to slowly cool to 25°C. Binding reactions were done in buffer containing 25 mM Hepes (pH 7.7), 40 mM KCl, 0.25 mM EDTA, 0.5 mM DTT, 10 μg/ml BSA, 10 mM $MgCl_2$ and 10% glycerol, in the presence of 5000 c.p.m of labelled DNA (~10 nM), and when indicated 1 μg of poly(dI-dC) and 0.5 and 1 μM of protein. The reactions were incubated at 25°C for 5 min and analysed on 5% native TBE PAGE. Gels were dried and analysed using a Fuji PhosphorImager.

XerCD/*dif* recombination assay

Recombination was measured as described in (21,29). Briefly, strains carrying the Δ(*dif*)$_{33}$ and *xerC::*Gm mutations, an insertion of the *dif-lacI-dif* cassette and a plasmid producing the γ or 3γ proteins were grown in L broth plus 0.025% arabinose, rendered competent and transformed with pFC242 (XerC). Transformants were placed in LB-agar containing 20 μg/ml spectinomycin plus 100 μg/ml ampicillin and 0.025% arabinose and grown overnight at 37°C. Five independent transformants were inoculated in the same medium, grown for 5 h, diluted and placed in L broth plus X-gal (40 μg/ml). The ratio of dark blue to total colonies was used to calculate the frequency of *lacI* loss per generation. The mean and standard deviation of the 5 independent measures is plotted in the figures.

Genome analyses

Genome sequences were extracted from the Genome Review database with the following genome accession numbers: AE003852_GR.1 (*V. cholerae* chromosome 1), AE005176_GR.1 (*L. lactis* IL1403), AE007317_GR.1 (*Streptococcus pneumoniae* R6), AE014074_GR.1 (*Streptocccus pyogenes* MGAS315), AL009126_GR.3 (*B. subtilis* 168), AL732656_GR.1 (*Streptococcus agalactiae* NEM316), AM406671_GR.1 (*L. lactis* spp. *Cremoris* MG1363) and U00096_GR.2 (*E. coli* MG1655).

In all species, the analyses were carried on the leading strand as it is the relevant strand for KOPS/SRS activity. Leading strands were defined as the DNA strand reported in Genbank files downstream of the replication origin up to *dif* and the reverse complement strand from *dif* to the origin. In *B. subtilis,* we used the PLR position instead of the origin because it is reported that the skew of SRS shifts in this region and not at the origin (31). Ori/PLR and *dif* positions were, respectively, 3 923 767 and 1 588 801, 1 and 1 564 104, 3 965 606 and 1 942 543 in *E. coli*, *V. cholerae* and *B. subtilis*. For *Streptococcaceae,* the origin is at position 1 and the position of *dif* is 1 238 253 (*L. lactis* MG1363) (32), 1 259 289 (*L. lactis* IL1403), 1 009 512 (*S. agalactiae* NEM316), 1 039 995 (*S. pneumoniae* R6) and 893 748 (*S. pyogenes* MGAS315). The experimentally defined region of FtsK activity around *dif* (21) includes positions 1438 to 1776 kb that represents roughly 7% of the *E. coli* chromosome. For the analyses, we used in all species, a region representing 7% of the genome centred on *dif* that we call in this paper *dif* region.

Over-representation. Motif count analyses were performed on the leading strand of each strain. Since KOPS have a degenerate nucleotide, we analysed 'motifs families' counts: for example, the motif GGGNAGGG is represented by the family GGGAAGGG, GGGCAGGG, GGGGAGGG and GGGTAGGG. To assess over-representation of motifs of a given length, the observed count of each motif was compared to the count expected in random sequences showing the same oligonucleotide composition. The significance of the difference between the counts was evaluated by calculating the associated *P*-value, which is the probability that the count of a given motif in a random sequence under a Markov model of order 2 (see Supplementary Material) is greater than the observed count for this motif. The *P*-value was obtained using a compound Poisson approximation of motif counts (35), which has been shown to be reliable even when the sequence length is relatively short (as is the case for analyses in the *dif* region).

Skew significance. We define the leading strand skew of a motif as the number of its occurrences on the leading strand of the replication fork over the total number of its occurrences, in the observed region. The statistical significance of the motifs leading strand skew was evaluated by calculating the associated *P*-value, which is the probability that the skew of a given motif in a random sequence (under a Markov model of order 1, see below) is greater than the observed skew. The *P*-value was obtained using a Gaussian approximation of motif counts (10).

Orders of the Markov models used for score calculation. We determined empirically the order of the Markov model to use for evaluation of the over-representation score. We compared the rank of the KOPS motif scores in *E. coli*, *B. subtilis* and *V. cholerae* (chromosomes 1) in all possible models (Supplementary Figure S6). We chose the lowest order model (this insures good sampling of the model parameters) that minimized the ranks (thus showing over-representation of the motifs). We retained a Markov model of order 2

for the over-representation score, which is the minimal model to take into account codon bias. For the skew score, we used a model of order 1, which is sufficient to take into account the G/C skew. All calculations were performed using the RMES software (User guide: Hoebeke, M. and Schbath, S. (2006), ''R'MES: Finding Exceptional Motifs'', version 3. http://genome.jouy.inra.fr/ssb/rmes) and custom Perl scripts that are available upon request.

RESULTS AND DISCUSSION

Covalent trimers of FtsKγ bind KOPS

Most FtsK homologues, including in *L. lactis*, contain a domain homologous to the *E. coli* FtsKγ subdomain in sequence and length located at their C-terminal end (Figure 1A), suggesting a conserved role in DNA binding and the control of translocation. We thus attempted to characterize KOPS motifs using the DNA binding activity of purified *E. coli* and *L. lactis* FtsKγ subdomains. However, the poor affinity of purified *E. coli* FtsKγ subdomain (γ_{Ec}) to DNA containing KOPS motifs, and the fact that it can be detected in EMSA experiments only in the absence of competitor DNA complicated this approach [(27), data not shown]. Since three γ_{Ec} monomers are involved in the interaction with a single KOPS in the γ_{Ec}/KOPS co-cristal structure (26), we reasoned that chimera protein containing three FtsKγ might bind KOPS with a higher affinity than FtsKγ monomers. We constructed a gene coding for a chimera protein, $3\gamma_{Ec}$, that contains three *E. coli* γ subdomains separated by linkers rich in glycine residues (Figure 1B), predicted to be flexible and already successfully used for the construction of covalent multimers of different FtsK domains (36). This protein was fused to His and FLAG tags at its N-terminal end for subsequent purification and western blot analysis (see 'Materials and Methods'; Supplementary Figure S1).

We assayed the functionality of the $3\gamma_{Ec}$ protein by testing the known activities of γ_{Ec}. The induction of XerCD/*dif* recombination was tested using a Δ(*lacI*) *E. coli* strain carrying a *dif-lacI-dif* recombination cassette inserted in place of the *dif* site on the chromosome, which allows accurate measurement of recombination frequencies (29). This strain was rendered Δ(*ftsKγ*), and the γ_{Ec} or $3\gamma_{Ec}$ protein were produced from a plasmid (see 'Materials and Methods'). As previously reported, *ftsKγ* deletion drastically reduced the recombination frequency compared to *ftsK*wt and γ_{Ec} production partially restored recombination (Figure 1C) (29). The $3\gamma_{Ec}$ protein was readily produced *in vivo* at levels comparable to FtsKγ_{Ec} alone in the same conditions (see 'Materials and Methods'; Supplementary Figure S1) and had the same activity as γ_{Ec} for the induction of XerCD/*dif* recombination (Figure 1C). These results are consistent with a recent report showing that covalent trimers of *E. coli* FtsKγ can induce XerCD/*dif* recombination *in vitro* and *in vivo* between plasmid-borne *dif* sites (37). We concluded that the $3\gamma_{EC}$ protein displays the same activity as FtsKγ_{Ec} for the induction of XerCD/*dif* recombination.

The $3\gamma_{Ec}$ protein was purified (see 'Materials and Methods'; Supplementary Figure S1), and its capacity to bind KOPS-containing DNA was assayed in EMSA experiments. We used different DNA fragments containing either one KOPS or three non-overlapping KOPS 6 bp apart and assayed binding at two $3\gamma_{Ec}$ concentrations. This was done in the presence or not of a large excess of competitor DNA devoid of KOPS motifs (polydIdC). The single KOPS-containing DNA was slightly shifted after incubation with $3\gamma_{Ec}$, the $3\gamma_{Ec}$-DNA complexes not migrating as a single shifted band but forming a smear immediately up to the free DNA (Figure 1D). This smear disappeared in the presence of competitor DNA. These results are reminiscent of the poor binding efficiency observed using a purified FtsKγ monomer and a DNA containing three overlapping KOPS (27). In contrast, binding of $3\gamma_{Ec}$ to a DNA containing three non-overlapping KOPS was clearly detectable and formed a major complex in the absence of competitor DNA (Figure 1E). This complex appeared unstable during migration but formed efficiently even in the presence of competitor DNA (Figure 1E, right). These results combined with previously reported data show that the FtsKγ-KOPS interaction is poorly efficient and forms unstable complexes that are dissociated during migration in EMSA experiments and are displaced by an excess of non-specific DNA. FtsKγ-KOPS complexes are nevertheless readily detected in EMSA experiments using the $3\gamma_{Ec}$ protein, which renders possible the characterization of KOPS from different species using this *in vitro* assay.

Lactococcus lactis FtsKγ does not recognize KOPS or AIMS motifs

We constructed a gene coding for a chimera protein, $3\gamma_{Ll}$, equivalent to $3\gamma_{Ec}$ but containing three copies of the FtsKγ_{Ll} subdomain (see 'Materials and Methods'). The $3\gamma_{Ll}$ protein was produced at quantities equivalent to $3\gamma_{Ec}$ (Supplementary Figure S1). $3\gamma_{Ll}$ was purified and we assayed binding to the DNA containing three *E. coli* KOPS (Figure 1F). No binding was detectable in conditions where the $3\gamma_{Ec}$-KOPS complexes are readily detected (compare with Figure 1E). This suggested that *L. lactis* FtsK does not recognize *E. coli* KOPS. This hypothesis was consistent with the distribution of *E. coli* KOPS on the *L. lactis* chromosome (Figure 2). Whereas KOPS are numerous and highly skewed on the *E. coli* chromosome, they are infrequent and poorly skewed on the *L. lactis* genome. From these criteria, the SRS motif appeared as even worse candidate than KOPS to fulfil the role of KOPS in *L. lactis* (Figure 2).

We next assayed the four AIMS motifs reported for *L. lactis*. (6). We assayed binding of the $3\gamma_{Ll}$ protein to four different DNA fragments, each containing three copies of a particular AIMS motif. Three of the four fragments yielded no detectable binding (Supplementary Figure S2). The fourth fragment, containing three consecutive 5′-AAGAAGAT-3′ motif, was reproducibly slightly shifted by $3\gamma_{Ll}$ (Supplementary Figure S2). This

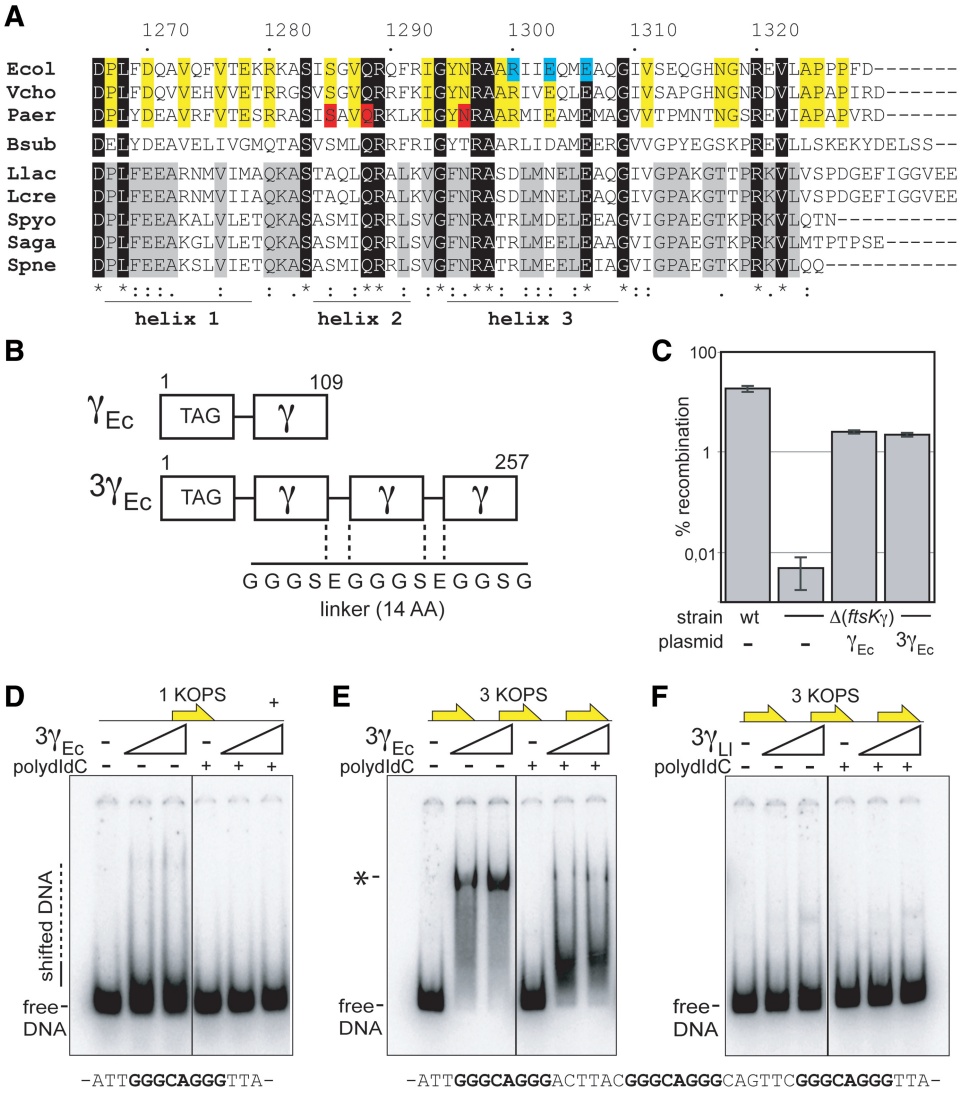

Figure 1. Chimerical protein $3\gamma_{Ec}$ efficiently binds KOPS-containing DNA. (**A**) Alignment of the winged-helix domains of different FtsKγ. Helices were positioned from the *E. coli* FtsKγ structure. Coordinates (AA) are those of *E. coli* FtsK. Invariant residues are highlighted in black and indicated by stars. Double dots indicate conserved and single dots semi-conserved residues. Residues common to *E. coli* (Ecol), *V. Cholerae* (Vcho) and *P. aeruginosa* (Paer), which share the same KOPS motifs, are highlighted in yellow. Residues common to *Streptococcaceae* are highlighted in grey. Residues of *E. coli* FtsK whose mutation to alanine results in a KOPS-blind phenotype are highlighted in blue. Residues possibly involved in specific base recognition in the *P. aeruginosa* FtsKγ/KOPS co-crystal are highlighted in red. Colon and dot indicate conserved and semi-conserved residues, respectively. Abbreviations: Bsub, *B. subtilis*; Llac, *L. lactis* subsp. *lactis*, Lcre, *L. lactis* subsp. *cremoris*; Spyo, *S. pyogenes*; Saga, *S. agalactiae*; Spne, *S. pneumoniae*. (**B**) Schematic representation of a covalent trimer. Three γ subdomains of *E. coli* FtsK (64AA) were connected by a flexible linker (16AA). The resulting 3γ protein ($3\gamma_{Ec}$) was His-Flag-tagged in the amino-terminal part (see 'Materials and Methods'). (**C**) Activation of XerCD/*dif* recombination by the *E. coli* 3γ protein. Recombination frequency were measured in Δ(*lacI xerC ftsKγ*) *E. coli* strains carrying the *dif-lacI-dif* recombination cassette in place of the *dif* site and a plasmid encoding the $FtsK\gamma_{EC}$ or $3\gamma_{EC}$ proteins shown in (B), as indicated. Recombination was induced by transformation with plasmid pFC241 (XerC) and measured as previously described (21,29). (**D–F**) EMSA performed with the indicated DNA substrate containing one KOPS motif (D) or three KOPS motifs (E and F) and the indicated protein: $3\gamma_{EC}$ (D and E) or $3\gamma_{Ll}$ (F). The relevant DNA sequences are shown below the gel. Reactions contained 0.5 and 1 μM of 3γ protein. The presence (+polydIdC) or absence of competitor (−polydIdC) or protein (−) is indicated. The free DNA probes and the shifted DNA are indicated. Asterisk indicates the major complex formed in the absence of competitor DNA.

binding activity, however, appeared largely weaker than binding of the $3\gamma_{Ec}$ protein to a DNA containing three KOPS (compare Supplementary Figure S2 with Figure 1E). This weak binding may be due to a faint activity of the $3\gamma_{Ll}$ protein compared to its *E. coli* counterpart. Alternatively, the *L. lactis* KOPS motif may differ from both known KOPS and AIMS motifs. To differentiate between these two hypotheses, we attempted to improve KOPS prediction and find better candidate motifs in the *L. lactis* genome.

Definition of prediction criteria for KOPS

We reasoned that the common properties of KOPS and SRS motifs distribution in their respective genomes should allow us to establish prediction criteria for KOPS in

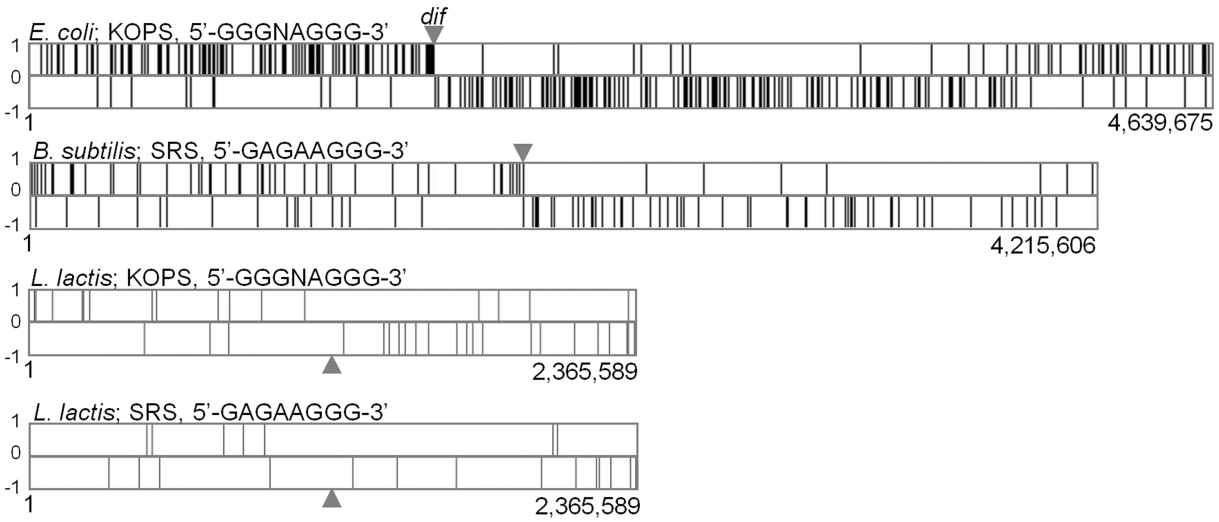

Figure 2. *Escherichia coli* KOPS and *B. subtilis* SRS motifs are bad candidate motifs for *L. lactis* KOPS motifs. The graphs show distribution of KOPS or SRS motifs in relevant bacterial genomes. Genomes and motifs are indicated. Coordinates are in bp. Grey arrowheads show the position of the chromosome dimer resolution site. The sequence is red on the top DNA strand; a +1 bar indicates a motif and a −1 bar its complementary sequence. Graphs were generated using an in-house version of the FindOligomers software (5).

Table 1. Properties of known KOPS motifs

Species (Motif)	Region analysed	Leading strand skew			Over-representation		
		Skew[a]	*P*-value[b]	Rank[c]	Frequency[d]	*P*-value[e]	Rank[f]
E. coli GGGNAGGG	Complete genome	0.91	7.72×10^{-24}	24	1/12.6 kb	7.66×10^{-16}	3693
	dif region	1	0	1	1/11.6 kb	8.15×10^{-7}	413
V. cholerae chr. 1 GGGNAGGG	Complete genome	0.80	3.85×10^{-5}	703	1/21.9 kb	7.14×10^{-12}	2648
	dif region	0.92	0.054	7239	1/15.9 kb	1.31×10^{-4}	388
B. subtilis GAGNAGGG	Complete genome	0.79	5.83×10^{-4}	5817	1/12.7 kb	1.47×10^{-11}	3621
	dif region	0.90	0.07	11 103	1/10.2 kb	8.84×10^{-4}	1312

[a]The skew is the ratio of number of motifs on the leading strand to total number of motifs.
[b]The *P*-value evaluates the probability that the observed skew is explained by chance.
[c]Skew rank: all motifs are ranked according to their skew significance: the lower the rank, the more significantly skewed the motif.
[d]$1/x$ kb correspond to the average frequency of motifs in the region of interest expressed as 1 motif per x kilo-base.
[e]The frequency *P*-value evaluates the probability that the observed frequency is explained by chance.
[f]Over-representation rank: all motifs are ranked according to their over-representation: the lower the rank, the more over-represented the motif.

L. lactis. Since both KOPS and SRS are octamers and the data available for motifs consensus show that the *E. coli* KOPS is degenerated at least at the fourth position, we analysed all families of octamers degenerated at one of any positions (see 'Materials and Methods'; Table 1). As previously shown (2,8), KOPS and SRS motifs are significantly over-represented and skewed, with more than 75% present on the leading strand (Table 1). Indeed, a combination of the over-representation and skew scores identified the *E. coli* KOPS motif as one of the five most exceptional motifs [(2); Supplementary Figure S3A]. However, the same criteria did not discriminate clearly enough the KOPS motif in *V. cholerae* and the SRS motif in *B. subtilis* from all other octamers (Supplementary Figure S3B and S3C), suggesting that additional criteria are necessary to *de novo* predict KOPS motifs.

As *E. coli* FtsK acts mainly in a ~350-kb region around *dif* that represents ~7% of its genome (21), we speculated that KOPS distribution might be particularly important in this region. We looked for specific properties of KOPS/SRS in the equivalent region (here called the *dif* region) in *V. cholerae* and *B. subtilis* genomes. The skew of KOPS and SRS motifs, already high on the whole genome, was even higher in the *dif* region where ~90% of them were on the leading strand (Table 1). They also show an increased frequency (higher than 1/16 kb) and are significantly over-represented in this region (Table 1). This suggested that criteria for prediction of KOPS motifs should include a minimal leading strand skew in the *dif* region as well as a minimal frequency in this region. We chose as selective criterion a minimal skew of 90% in the *dif* region, because this is the most important property of KOPS and SRS motifs with respect to their activity. The minimal frequency was set conservatively at 1 motif every 40 kb because frequency is less critical to KOPS activity. Analyzing all octamers in *E. coli*, *V. cholerae*

and *B. subtilis* with these additional criteria allowed us to identify the KOPS and SRS motifs among the best candidates (Supplementary Figure S3D–S3F, respectively, where red dots respond to the *dif* region criteria). In addition, KOPS and SRS motifs were found among the most skewed motifs in the *dif* regions of the three genomes (reported as the intensity of red dots in Supplementary Figure S3D–S3F). We thus defined the following criteria for KOPS prediction: (i) a high skew and over-representation score on the whole genome, (ii) an occurrence of at least 1 every 40 kb in the *dif* region and (iii) a skew of at least 90% in the *dif* region with higher attention paid to the most skewed motifs.

Prediction of KOPS candidates in *L. lactis*

We used the criteria defined above to predicted possible KOPS motifs in *L. lactis*. As KOPS and SRS motifs are octamers, we initially analysed the distribution of all octamers degenerated at one position on the leading strand of the *L. lactis* subsp. *lactis* IL1403 genome (see 'Materials and Methods'; Figure 3A). The twenty best octamer candidates from our prediction criteria tended to be rich in purine bases (Supplementary Table 1). Interestingly, we noticed that 7 out of these 20 candidate motifs were composed of three very similar heptamer sub-motifs: 5′-GNAGAAG-3′, 5′-GANGAAG-3′ or 5′-GAAGNAG-3′ (Supplementary Table 1). This suggested that the *L. lactis* KOPS could be a heptamer. Indeed, when applying the same prediction criteria to all heptamers in this species, the motif 5′-GAAGAAG-3′ had a particularly striking distribution (Figure 3B): this motif is very frequent (1/2.1 kb), as it corresponds to the 10th most over-represented motif on the leading strand (P-value = 2.15×10^{-160}) and the fifth on the *dif* region. Its skew is also very high on the whole genome and >91% in the *dif* region.

To check if the properties of the 5′-GAAGAAG-3′ motif was conserved in bacteria related to *L. lactis* subsp. *lactis*, we first analysed the genome of *L. lactis* subsp. *cremoris* MG1363, which displays an average of 15% DNA divergence with the *L. lactis* subsp. *lactis* genome (38), although its FtsKγ subdomain is strictly identical (Figure 1A). The 5′-GAAGAAG-3′ motif was the best KOPS candidate motif in *L. lactis* subsp. *cremoris*, strengthening our assumption that it might function as KOPS in *L. lactis*. We then analysed skewed heptamers in the genomes of other *Streptococcaceae*: *S. pneumoniae*, *S. agalactiae* and *S. pyogenes* (Supplementary Figure S5). These bacteria harbour similar FtsKγ subdomains that diverge from the *L. lactis* FtsKγ (Figure 1A). In these species, the 5′-GAAGAAG-3′ motif, although over-represented, was not skewed enough to fulfil our criteria. Furthermore, the sequence of the best candidate motifs varied between these three species: 5′-GC AGATG-3′ in *S. pneumoniae*, 5′-GAAGCAG-3′ in *S. agalactiae* and 5′-GTAGAAG-3′ in *S. pyogenes* (Supplementary Figure S5 and Supplementary Table S2). These motifs show a sequence related to but different from the 5′-GAAGAAG-3′ motif. This suggests that the KOPS

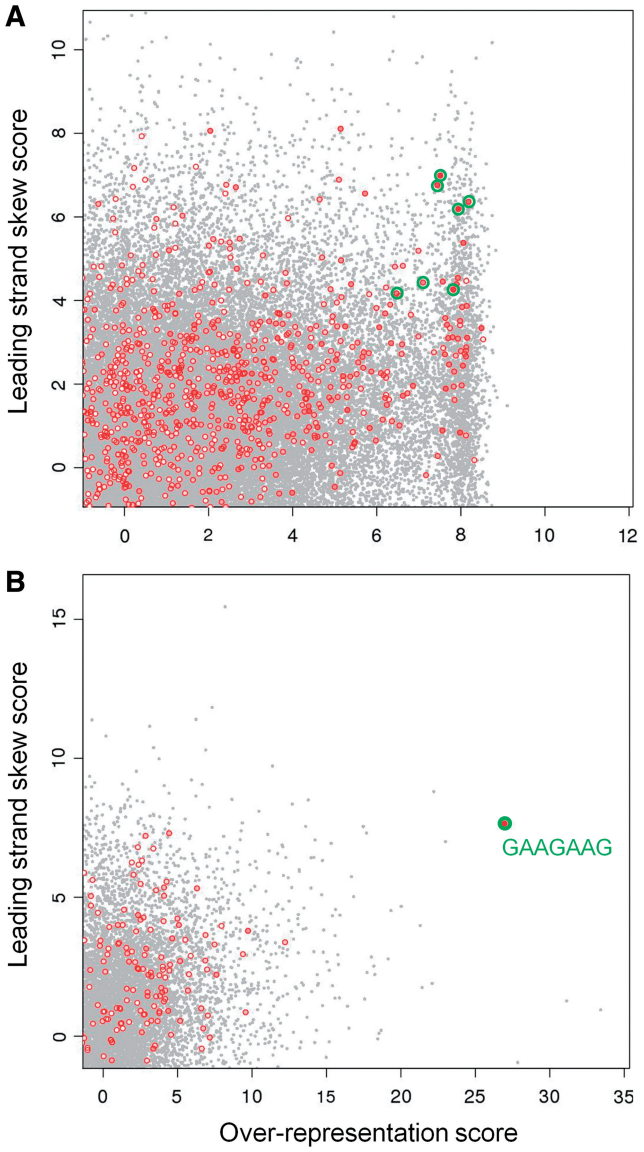

Figure 3. Prediction of the *L. lactis* KOPS motif. (**A**) Distribution of octamers with one degenerated position in the *L. lactis* chromosome. All octamers that have a positive over-representation and skew score and a minimal frequency of 1 every 70 kb on the whole chromosome are represented in grey. Among these, motifs that have a specific distribution in the *dif* region are shown in red. They represent the 100th most skewed motifs among those with a leading strand skew in the *dif* region higher than 90%, a minimal frequency of 1 occurrence every 40 kb in this region and a minimal frequency of 1 occurrence every 70 kb in the whole genome. The strength of the red represents the over-representation score in the *dif* region (the stronger the red, the more over-represented the motif). Four motifs stand out as potential KOPS candidates (highlighted in green) with three additional ones that are slightly less exceptional but share very similar sequence (also in green). (**B**) Distribution of heptamers in the *L. lactis* chromosome. Same criteria and color code as in (A). One motif, 5′-GAAGAAG-3′ clearly stands out (shown circled in green).

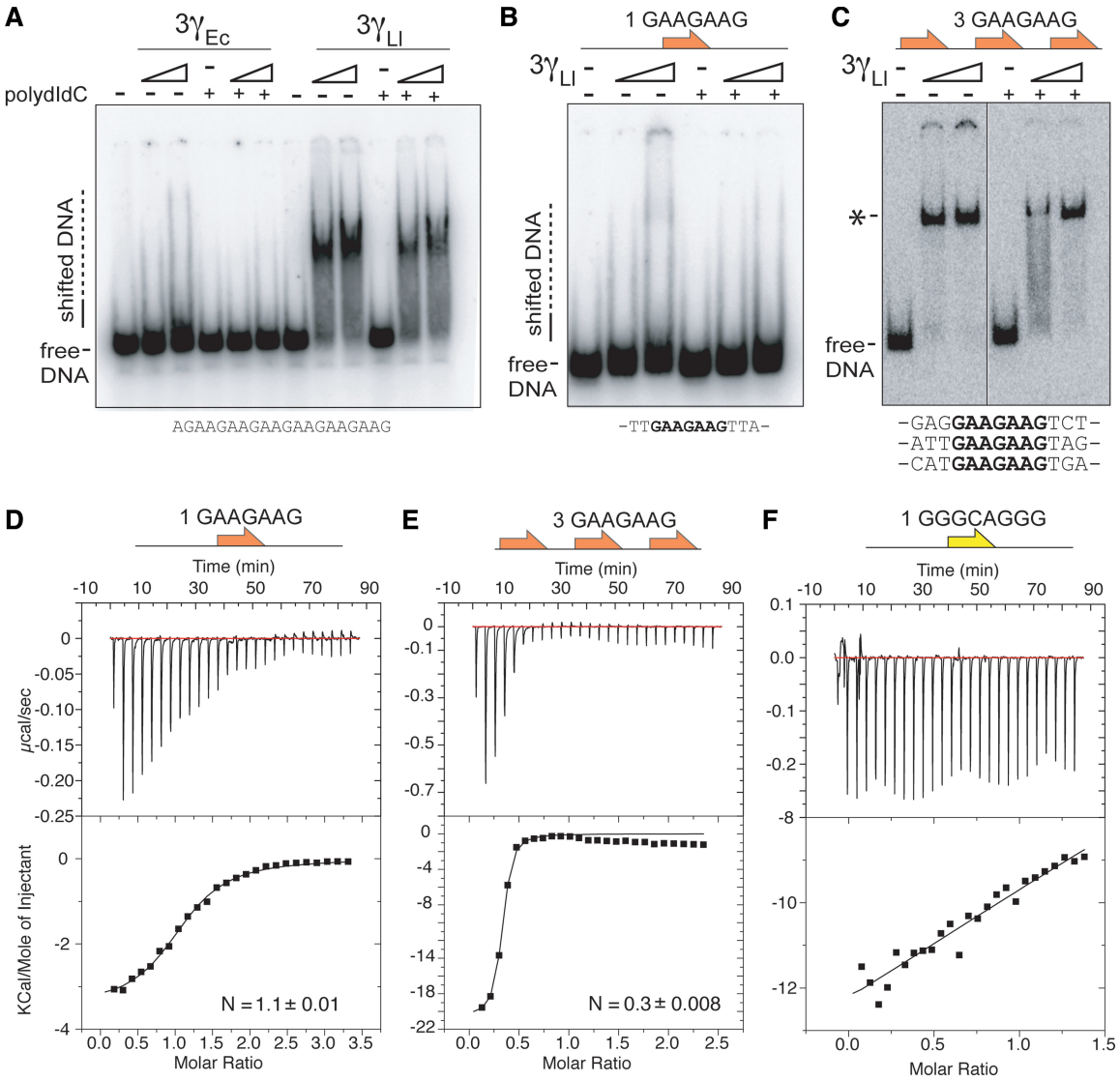

Figure 4. The $3\gamma_{Ll}$ protein recognizes the 5′-GAAGAAG-3′ heptamer. (**A–C**) Same EMSA experiment as in Figure 1D–F, with DNA substrates containing a repetition of 5′-GAA-3′ motifs (A), a single 5′-GAAGAAG-3′ motif (B), or three non-overlapping 5′-GAAGAAG-3′ motifs (C) and the indicated proteins. The relevant DNA sequences are shown below the gels. **(D–F)** ITC experiments performed by titrating the $3\gamma_{Ll}$ protein with the indicated DNA substrates: (D) containing a single 5′-GAAGAAG-3′ motif, (E) three non-overlapping 5′-GAAGAAG-3′ motifs and (F) a single *E. coli* KOPS (5′-GGGCAGGG-3′). The best-fitted model curves using a 'one set of sites' model are shown (continuous line) with the corresponding stoichiometry values indicated (N).

motif might be less conserved in *Streptococcaceae* compared to γ-proteobacteria.

Lactococcus lactis FtsKγ binds the 7 bp 5′-GAAGAAG-3′ motif

The above analysis showed that KOPS candidate motifs in *L. lactis* were 7 or 8 bp poly-purine tracks, a number of them containing the GAA trinucleotide. We therefore first assayed a DNA containing six consecutive GAA for $3\gamma_{Ll}$ binding in EMSA experiments. This DNA was shifted efficiently by $3\gamma_{Ll}$ even in the presence of competitor DNA (Figure 4A). In contrast, $3\gamma_{Ec}$ barely bound the DNA containing consecutive GAA and only in the absence of competitor DNA. Thus, a DNA containing consecutive GAA trinucleotides is efficiently and specifically bound by $3\gamma_{Ll}$. We then directly assayed the 5′-GAAGAAG-3′ heptamer, which is the best candidate found in our predictive approach (Figure 3). $3\gamma_{Ll}$ bound a DNA fragment containing a single 5′-GAAGAAG-3′ motif poorly in a similar manner that $3\gamma_{Ec}$ bound a fragment containing a single KOPS (compare Figure 4B with 1D). We thus constructed a DNA containing three non-overlapping 5′-GAAGAAG-3′ motifs separated by 6 bp. $3\gamma_{Ll}$ formed specific complexes with this DNA both in the absence and presence of competitor DNA (Figure 4C). We concluded that the 5′-GAAGAAG-3′ heptamer is sufficient for specific binding by the $3\gamma_{Ll}$ protein. However, since KOPS and SRS motifs are 8 bp long, we considered the possibility that octamers may be better substrates than the 5′-GAAGAAG-3′ heptamer. We thus assayed

the octamers contained in the GAA polymer used in Figure 4A. Of these, the 5′-AAGAAGAA-3′ motif was not recognized whereas the 5′-AGAAGAAG-3′ and 5′-GAAGAAGA-3′ motifs were slightly shifted by $3\gamma_{LI}$ in a manner that resembled binding to a single 5′-GAAGAAG-3′ motif (compare Supplementary Figure S4 with Figure 4B). These results show that (i) the preferred $3\gamma_{LI}$ binding sequence is longer than 6 bp since all possible 6 bp or shorter motifs are contained into the 5′-AAGAAGAA-3′ motif that is not recognized by $3\gamma_{LI}$; (ii) the $3\gamma_{LI}$ binding sequence is the 5′-GAAGAAG-3′ heptamer since the three possible octamers containing this motif and contained into the 5′-GAA-3′ concatemer shown in Figure 4 are not better binding sites. Both the sequence and the shorter length of the $3\gamma_{LI}$ binding motif compared to *P. aeruginosas* and *E. coli* KOPS suggest a different mode of binding. Indeed, in the *P. aeruginosas* FtsKγ/KOPS complex, the three FtsKγ monomers are located head to tail. Two monomers recognize the two GGG repeats of the KOPS motif, while the third appears to stabilize the complexes by protein–protein interaction and may recognize the central NA (26). The shorter length of the *L. lactis* motif would imply a different geometry for FtsKγ monomers arrangement in the complex. In addition, the absence of direct repetition of base triplex at the edges of the 5′-GAAGAAG-3′ motif appears inconsistent with the mode of binding described for *P. aeruginosas*.

Since the interaction of $3\gamma_{LI}$ with a single 5′-GAAGAAG-3′ motif appears unstable and poorly specific in EMSA experiments, we measured the $3\gamma_{LI}$/5′-GAAGAAG-3′ interaction using isothermal titration calorimetry (ITC; see 'Materials and Methods'). Results showed that $3\gamma_{LI}$ formed stable complexes with DNA containing either one or three 5′-GAAGAAG-3′ motifs (Figure 4D and E). The patterns obtained fitted well with a DNA/protein stoichiometry 1:1 and 3:1 for DNA containing one and 3 5′-GAAGAAG-3′ motifs, respectively. This indicates that one molecule of $3\gamma_{LI}$ binds to one 5′-GAAGAAG-3′ motif (therefore, three molecules of $3\gamma_{LI}$ associate with one DNA molecule containing three motifs). In contrast, the ITC data resulting from the titration of $3\gamma_{LI}$ by a DNA containing one *E. coli* KOPS motif failed to indicate any binding (Figure 4). Taken together, these data strongly suggest that the γ subdomain of *L. lactis* FtsK specifically recognizes the 5′-GAAGAAG-3′ motif.

The 5′-GAAGAAG-3′ motif controls FtsK translocation *in vivo*

To assay the activity of the 5′-GAAGAAG-3′ motif *in vivo*, we took advantage of the role of FtsK in the induction of XerCD/*dif* recombination. We have previously reported that inserting three consecutive KOPS in non-permissive orientation next to a *dif* site lowers its capacity to recombine in *E. coli* (29). Non-permissive KOPS are thought to promote FtsK loading and subsequent translocation away from *dif* thereby lowering its capacity to reach the XerCD/*dif* complex (26). We also previously constructed an *E. coli* strain carrying the

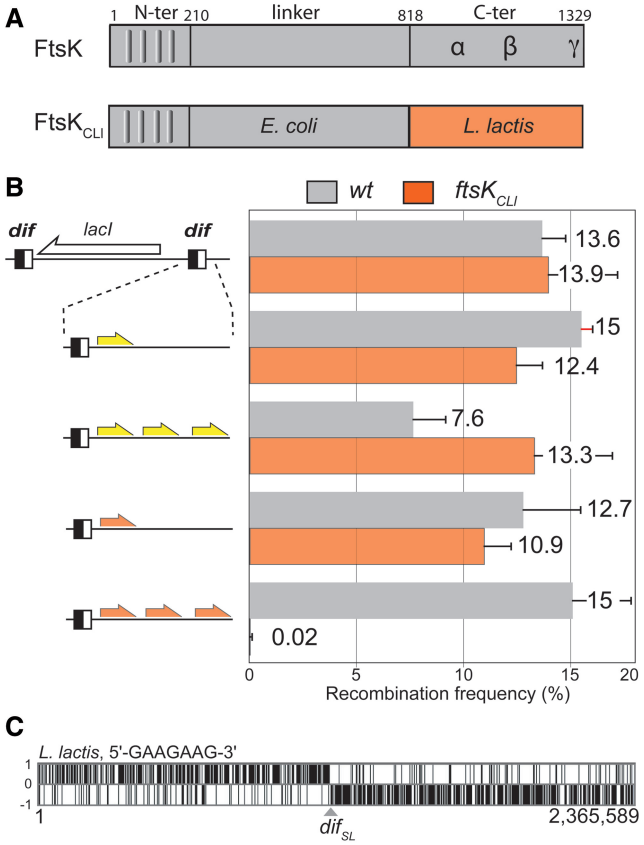

Figure 5. The 5′-GAAGAAG-3′ motif controls *L. lactis* FtsK activity *in vivo*. (**A**) Structure of the two FtsK proteins used. Top: wild-type *E. coli* FtsK with domains and subdomains indicated (grey). Vertical bars represent the transmembrane segments in the N-terminal domain. Coordinates are in AA. Bottom: the C-terminal domain of FtsK has been replaced by its homolog from *L. lactis* (red) yielding the FtsK$_{CLI}$ chimeral protein. (**B**) Measure of the recombination frequencies between *dif* sites inserted in direct repetition at the *dif* position on the *E. coli* chromosome. The relevant structure of the *dif-lacI-dif* cassette and its derivative after insertion of KOPS motifs is shown. Yellow arrows: *Escherichia coli* KOPS motifs (5′-GGGCAGGG-3′); red arrows: 5′-GAAGAAG-3′ motifs. Consecutive motifs were separated by 6 bp of random DNA (see Figure 1E and 4C). Bars show means of five independent measures (shown right of the bars) with standard deviations. Frequencies are in percent per cell per generation. Grey bars: strains producing wt FtsK; Red bars: strains producing the FtsK$_{CLI}$ protein. (**C**) Distribution of the 5′-GAAGAAG-3′ motif on the *L. lactis* chromosome. The graphs were obtained as in Figure 2. Coordinates are in base pair. Grey arrowheads show the position of the chromosome dimer resolution site.

C-terminal part of *L. lactis ftsK* in place of its *E. coli* counterpart (Figure 5A). The resulting strain fully supported resolution of chromosome dimers, making this strain a useful tool to study FtsK$_{CLI}$ activities in a cellular context (29). To assay the effect of the 5′-GAAGAAG-3′ motif on FtsK activity, we constructed a set of strains carrying either the *E. coli ftsK* or *ftsK$_{CLI}$* gene and non-permissive *E. coli* KOPS (5′-GGGCAGGG-3′) or 5′-GAAGAAG-3′ motifs next to a *dif* site of a *dif-lacI-dif* construct inserted in place of the *dif* site (Figure 5B). Recombination was scored as the appearance of dark blue colonies on indicator medium-containing plates [see 'Materials and Methods'; (29)]. As previously

shown, insertion of three non-permissive KOPS lowered *E. coli* FtsK-driven recombination about 2-fold (Figure 5B). A single non-permissive *Escherichia coli* KOPS yielded no significant effect in this assay. No significant effect of non-permissive 5′-GAAGAAG-3′ motif was detected on *E. coli* FtsK-driven recombination, showing that this motif has no KOPS activity on *E. coli* FtsK. On the other hand, *E. coli* KOPS had no effect on $FtsK_{CLl}$-driven recombination, showing that *E. coli* KOPS do not control *L. lactis* FtsK translocation (Figure 5). As in the *E. coli* FtsK/KOPS system, a single non-permissive 5′-GAAGAAG-3′ motif had no significant effect on $FtsK_{CLl}$-driven recombination frequencies. However, three consecutive 5′-GAAGAAG-3′ motifs lowered recombination frequencies almost 100 times (Figure 5, last line, compare with the two time effect yielded by *E. coli* KOPS). This high level of inhibition might reflect a higher efficiency of *L. lactis* KOPS compared to *E. coli* KOPS and/or a lower activity of the $FtsK_{CLl}$ protein compared to *E. coli* FtsK. These results show that the 5′-GAAGAA G-3′ motif controls *L. lactis* FtsK translocation in a cellular context.

CONCLUSION

Using a combination of predictive and functional approaches, we have successfully characterized a DNA motif showing KOPS activity on the *L. lactis* FtsK translocase. This shows that bacteria into which known KOPS motifs are infrequent and/or poorly skewed, as the *Streptococcaceae* family, nevertheless use other KOPS motifs as chromosome segregation guides. KOPS-guided chromosome segregation is thus widely conserved in bacteria. The *L. lactis* KOPS motif differs both in sequence and in length from previously described KOPS and SRS motifs. FtsK homologues thus either have acquired independently different DNA binding specificities or changed specificity during evolution. To do so, a frequent and skewed motif has first to be selected. This motif may be only slightly skewed and/or not over-represented at first and then selected for higher skew and frequency due to its KOPS activity. Alternatively, the motif selected may be highly skewed and/or over-represented for reasons other than its KOPS activity. We assume this second hypothesis plausible because bacterial chromosomes usually contain numerous motifs showing remarkable skews and/or representations. For instance, the 5′-GAAGAAGA-3′ octamer, which contains the *L. lactis* KOPS, is extremely over-represented in bacterial genomes from nearly all phylogenetic groups (39). The fact that the *L. lactis* KOPS motif is A-rich when the *L. lactis* genome is AT-rich also argues towards this hypothesis.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online: Supplementary Tables 1 and 2, and Supplementary Figures 1–6.

ACKNOWLEDGEMENTS

We thank Yves Quentin for critical reading, Laurent Paquereau for helpful discussion and Virginie Gervais for technical assistance on the MicroCal ITC200 Isothermal Titration Calorimeter provided by the PICT(IBISA) platform in Toulouse. We are grateful to the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources.

FUNDING

CNRS, University Paul Sabatier, ANR contract (BLAN06-2 134012 to F.C. and ANR-07-BLAN-0367-03 to M.E.K.); a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche and the Fondation pour la Recherche Médicale (to S.K.). Funding for open access charge: ANR contract BLAN06-2 134012 and ANR contract BLAN0367-03.

Conflict of interest statement. None declared.

REFERENCES

1. Rocha,E.P. (2008) The organization of the bacterial genome. *Annu. Rev. Genet.*, **42**, 211–233.
2. Touzain,F., Petit,M.A., Schbath,S. and El Karoui,M. (2011) DNA motifs that sculpt the bacterial chromosome. *Nat. Rev. Microbiol.*, **9**, 15–26.
3. Blattner,F., Plunkett,G.r., Bloch,C., Perna,N., Burland,V., Riley,M., Collado-Vides,J., Glasner,J., Rode,C., Mayhew,G. *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1462.
4. Grigoriev,A. (1998) Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.*, **26**, 2286–2290.
5. Salzberg,S.L., Salzberg,A.J., Kerlavage,A.R. and Tomb,J.F. (1998) Skewed oligomers and origins of replication. *Gene*, **217**, 57–67.
6. Hendrickson,H. and Lawrence,J.G. (2006) Selection for chromosome architecture in bacteria. *J. Mol. Evol.*, **62**, 615–629.
7. Dillingham,M.S. and Kowalczykowski,S.C. (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.*, **72**, 642–671.
8. Bigot,S., Saleh,O., Lesterlin,C., Pages,C., El Karoui,M., Dennis,C., Grigoriev,M., Allemand,J., Barre,F. and Cornet,F. (2005) KOPS: DNA motifs that control E. coli chromosome segregation by orienting the FtsK translocase. *EMBO J.*, **24**, 3770–3780.
9. Levy,O., Ptacin,J., Pease,P., Gore,J., Eisen,M., Bustamante,C. and Cozzarelli,N. (2005) Identification of oligonucleotide sequences that direct the movement of the Escherichia coli FtsK translocase. *Proc. Natl Acad. Sci. USA*, **102**, 17618–17623.
10. Halpern,D., Chiapello,H., Schbath,S., Robin,S., Hennequet-Antier,C., Gruss,A. and El Karoui,M. (2007) Identification of DNA motifs implicated in maintenance of bacterial core genomes by predictive modeling. *PLoS Genet.*, **3**, 1614–1621.
11. Amundsen,S.K., Taylor,A.F., Chaudhury,A.M. and Smith,G.R. (1986) recD: the gene for an essential third subunit of exonuclease V. *Proc. Natl Acad. Sci. USA*, **83**, 5558–5562.
12. Kooistra,J., Haijema,B.J. and Venema,G. (1993) The *Bacillus subtilis* addAB genes are fully functional in *Escherichia coli*. *Mol. Microbiol.*, **7**, 915–923.
13. el Karoui,M., Ehrlich,D. and Gruss,A. (1998) Identification of the lactococcal exonuclease/recombinase and its modulation by the putative chi sequence. *Proc. Natl Acad. Sci. USA*, **95**, 626–631.
14. Chédin,F., Noirot,P., Biaudet,V. and Ehrlich,S.D. (1998) A five-nucleotide sequence protects DNA from exonucleolytic degradation by AddAB, the RecBCD analogue of *Bacillus subtilis*. *Mol. Microbiol.*, **29**, 1369–1377.

15. Sourice,S., Biaudet,V., El Karoui,M., Ehrlich,S.D. and Gruss,A. (1998) Identification of the chi site of *Haemophilus influenzae* as several sequences related to the *Escherichia coli* chi site. *Mol. Microbiol.*, **27**, 1021–1029.
16. Bigot,S., Sivanathan,V., Possoz,C., Barre,F. and Cornet,F. (2007) FtsK, a literate chromosome segregation machine. *Mol. Microbiol.*, **64**, 1434–1441.
17. Iyer,L.M., Makarova,K.S., Koonin,E.V. and Aravind,L. (2004) Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. *Nucleic Acids Res.*, **32**, 5260–5279.
18. Sherratt,D.J., Arciszewska,L.K., Crozat,E., Graham,J.E. and Grainge,I. (2010) The *Escherichia coli* DNA translocase FtsK. *Biochem. Soc. Trans.*, **38**, 395–398.
19. Aussel,L., Barre,F., Aroyo,M., Stasiak,A., Stasiak,A. and Sherratt,D. (2002) FtsK Is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell*, **108**, 195–205.
20. Massey,T., Mercogliano,C., Yates,J., Sherratt,D. and Löwe,J. (2006) Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. *Mol. Cell*, **23**, 457–469.
21. Deghorain,M., Pagès,C., Meile,J.C., Stouf,M., Capiaux,H., Mercier,R., Lesterlin,C., Hallet,B. and Cornet,F. (2011) A defined terminal region of the *E. coli* chromosome shows late segregation and high FtsK activity. *Plos One*, **6**, e22164.
22. Espeli,O., Lee,C. and Marians,K. (2003) A physical and functional interaction between *Escherichia coli* FtsK and topoisomerase IV. *J. Biol. Chem.*, **278**, 44639–44644.
23. Steiner,W., Liu,G., Donachie,W.D. and Kuempel,P. (1999) The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers. *Mol. Microbiol.*, **31**, 579–583.
24. Graham,J.E., Sherratt,D.J. and Szczelkun,M.D. (2010) Sequence-specific assembly of FtsK hexamers establishes directional translocation on DNA. *Proc. Natl Acad. Sci. USA*, **107**, 20263–20268.
25. Bigot,S., Saleh,O., Cornet,F., Allemand,J. and Barre,F. (2006) Oriented loading of FtsK on KOPS. *Nat. Struct. Mol. Biol.*, **13**, 1026–1028.
26. Löwe,J., Ellonen,A., Allen,M., Atkinson,C., Sherratt,D. and Grainge,I. (2008) Molecular mechanism of sequence-directed DNA loading and translocation by FtsK. *Mol. Cell*, **31**, 498–509.
27. Sivanathan,V., Allen,M., de Bekker,C., Baker,R., Arciszewska,L., Freund,S., Bycroft,M., Löwe,J. and Sherratt,D. (2006) The FtsK gamma domain directs oriented DNA translocation by interacting with KOPS. *Nat. Struct. Mol. Biol.*, **13**, 965–972.
28. Yates,J., Zhekov,I., Baker,R., Eklund,B., Sherratt,D. and Arciszewska,L. (2006) Dissection of a functional interaction between the DNA translocase, FtsK, and the XerD recombinase. *Mol. Microbiol.*, **59**, 1754–1766.
29. Nolivos,S., Pages,C., Rousseau,P., Le Bourgeois,P. and Cornet,F. (2010) Are two better than one? Analysis of an FtsK/Xer recombination system that uses a single recombinase. *Nucleic Acids Res.*, **38**, 6477–6489.
30. Val,M., Kennedy,S., El Karoui,M., Bonné,L., Chevalier,F. and Barre,F. (2008) FtsK-dependent dimer resolution on multiple chromosomes in the pathogen *Vibrio cholerae*. *PLoS Genet.*, **4**, e1000201.
31. Ptacin,J., Nollmann,M., Becker,E., Cozzarelli,N., Pogliano,K. and Bustamante,C. (2008) Sequence-directed DNA export guides chromosome translocation during sporulation in Bacillus subtilis. *Nat. Struct. Mol. Biol.*, **15**, 485–493.
32. Le Bourgeois,P., Bugarel,M., Campo,N., Daveran-Mingot,M., Labonté,J., Lanfranchi,D., Lautier,T., Pagès,C. and Ritzenthaler,P. (2007) The unconventional Xer recombination machinery of *Streptococci/Lactococci*. *PLoS Genet.*, **3**, e117.
33. Cornet,F., Mortier,I., Patte,J. and Louarn,J. (1994) Plasmid pSC101 harbors a recombination site, psi, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site dif. *J. Bacteriol.*, **176**, 3188–3195.
34. Churchward,G., Belin,D. and Nagamine,Y. (1984) A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene*, **31**, 165–171.
35. Roquain,E. and Schbath,S. (2007) Improved compound Poisson approximation for the number of occurrences of any rare word family in a stationary Markov chain. *Advances in Applied Probability*, Vol. 39. Duke University Press, pp. 128–140.
36. Crozat,E., Meglio,A., Allemand,J., Chivers,C., Howarth,M., Vénien-Bryan,C., Grainge,I. and Sherratt,D. (2010) Separating speed and ability to displace roadblocks during DNA translocation by FtsK. *EMBO J*, **29**, 1423–1433.
37. Grainge,I., Lesterlin,C. and Sherratt,D.J. (2011) Activation of XerCD-dif recombination by the FtsK DNA translocase. *Nucleic Acids Res.*, **39**, 5140–5148.
38. Wegmann,U., O'Connell-Motherway,M., Zomer,A., Buist,G., Shearman,C., Canchaya,C., Ventura,M., Goesmann,A., Gasson,M.J., Kuipers,O.P. *et al.* (2007) Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. cremoris MG1363. *J. Bacteriol.*, **189**, 3256–3270.
39. Davenport,C.F. and Tümmler,B. (2010) Abundant oligonucleotides common to most bacteria. *PLoS One*, **5**, e9841.